CHAPTER 7
RESULTS AND DISCUSSION

This chapter explores the programming environment, data set, evaluation parameters of the proposed e-voting system with tabular and graphical result analysis. The proposed e-voting system is compared with various existing mechanism in order to display the comparative result justifications. The following performance measures Classification Accuracy (CA), Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are computed.

7.1 IMPLEMENTATION ENVIRONMENT

The experimental work is conducted on Intel Dual Core Processor, 2GB RAM running on windows7 ultimate with 500 GB memory. The proposed e-voting approach is implemented in NetBeans 8.0.2 IDE and JDK 1.8 programming environment. To evaluate the proposed algorithm with existing approach, the various existing classifier in Weka 3.7 library such as Naive Bayes, k-Nearest Neighbour (k-NN) and Support Vector Machine (SVM) are used.

7.1.1 Dataset

The MP Elections dataset is collected from various internet sources and government websites which are evaluated with Indian MP (Member of
Parliament) Election 2014 data sets state wise. It contains 35 records, attributes like state, total voters, female voters, male voters, total male turnout, total female turnout, % of male turnout, % of female turn out, total % turnout, total literacy, male literacy, female literacy, literacy growth state wise, total computer user, with internet computer user, without internet users and total internet user growth state wise.

7.2 EVALUATION PARAMETERS

In this section, the proposed algorithm is evaluated using classification accuracy and misclassification error. The approach suggests ECI (Election Commission India) to adapt e-voting system which increases the voter’s turnout, saves the travelling time and cost. Also, the system can predict internet user growth and literacy growth which indicates that increment of literacy and internet user growth are positive indication to adopt e-voting. It expresses the following performance matrix separately as CA, MAE and RMSE for the given attributes and instances. The definitions of these parameters are given in the following sub-sections.

7.2.1 Classification Accuracy

In text mining, the most common performance evaluation technique is classification accuracy. The classifier evaluation result is better if it has higher classification accuracy. It is defined by
7.2.2 Mean Absolute Error (MAE)

MAE is calculated as correctly classified MP election data behalf of overall data.

\[
MAE = \frac{1}{N} \sum_{i=1}^{n} |f_i - y_i|
\]

The mean absolute error is an average of the absolute errors where \( f_i \) is the prediction and \( y_i \) the true value.

7.2.3 Root Mean Squared Error (RMSE)

RMSE describes a mathematical model to evaluate the differences between the values predicted by a model and the values observed from the environment that is being modelled. These individual differences are also called residuals. The RMSE is calculated as:

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - y_i)^2}
\]

where \( \hat{y} \) is a vector of \( n \) predictions and \( y \) is the vector of observed values corresponding to the inputs to the function which generated the predictions.
7.3 EVALUATION PARAMETERS

The performance of proposed approach is evaluated in Java programming environment using Weka 3.7 library. For classification, 10 cross fold validation test is conducted to measure the precision, recall and F1 score, CA, MAE and RMSE for the following classifier namely Naive Bayes, k-NN and SVM with proposed algorithms. The results are obtained for MP Election 2014 datasets (State Wise) which are given in Table 7.1. It displays their average values for the respective parameter with respective datasets. The approach can classify internet user growth and literacy growth because the improvement in the internet user and literacy is a strong factor to adopt e-voting.

It expresses MAE and RMSE for given attribute and instances. The data giving details about the voter turnout through polling place voting in the parliamentary elections of India from 1952 to 2014 is not more than 66%. But through the proposed algorithm, it is proved that it can be raised up to 80% if e-voting is adopted. The following Figures 7.1 and 7.2 show the comparison of accuracy, MAE and RMSE values of existing algorithms and the proposed algorithm.
Table 7.1 Performance measures of the proposed approach using MP Election 2014 dataset

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>MAE</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naive Bayes</td>
<td>31.43</td>
<td>0.486</td>
<td>0.314</td>
<td>0.363</td>
<td>0.61</td>
<td>0.73</td>
</tr>
<tr>
<td>k-NN</td>
<td>48.57</td>
<td>0.556</td>
<td>0.486</td>
<td>0.517</td>
<td>0.51</td>
<td>0.70</td>
</tr>
<tr>
<td>SVM</td>
<td>77.14</td>
<td>0.7714</td>
<td>1.00</td>
<td>0.8709</td>
<td>0.23</td>
<td>0.47</td>
</tr>
<tr>
<td>Efficient Decision Tree</td>
<td>80.0</td>
<td>0.79411</td>
<td>1.0</td>
<td>0.8852</td>
<td>0.24</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Figure 7.1 Comparison of error values
According to the performance in Figure 7.1 and 7.2, the accuracy of proposed algorithm is 80% which is superior to Naive Bayes, k-NN and SVM for MP Election data set. Also, it is noticed that SVM is the closest technique to proposed algorithm compared to others. However, SVM is a functional classifier where it fails to reduce the misclassification error. In terms of CA, SVM tries to compete with the proposed algorithm. However, it fails to improve the accuracy because it is unable to remove irrelevant attributes. The proposed algorithm improves the accuracy with 2.86% than SVM.

**Figure 7.2 Accuracy of different algorithms**
This chapter explains about programming environment, data set, parameters evaluation and comparative result analysis of proposed methodology with existing methods. Initially, this chapter gives the programming environment with deployment processes like required hardware details and software details (Programming languages, Application Server, Storage database details) with required library descriptions. Next, it explains the required evaluations parameters like CA, MAE and RMSE for evaluating the performance of the proposed algorithm. The performance analysis is conducted with closest existing methods with respective parameters and given data set. Finally, it is noticed that the proposed algorithm is the best approach compared to other methods based on the results obtained.